

ESL GO! Newsletter

Special points of interest:

- End-of-year potluck party
- ESL Speaker Series: Mei-Hung Lin
- OEPT Rater Training
- Facebook
- Summer

End-of-year potluck!

Members of the Second Language Studies/ESL community at Purdue gathered Saturday, April 27 to relax and celebrate the end of the 2012-2013 academic year. A wide variety of dishes were brought to the potluck party. Students and faculty then simultaneously played some soccer and Frisbee, or FrisbeeSoccer as it is now known!



Potluck attendees on April 27, 2013.

ESL Speaker Series: A Corpus-based Approach to Identity Construction of L1 and L2 Writers in Academic Discourse

Mei-Hung Lin presented on L1 and L2 writers' identity as it is projected through science and humanities research articles. An analysis of linguistic features was conducted using a corpus-based approach. The presenter also explained that interviews were conducted in

order to determine how L1 and L2 writers construct the self through discourse. A fascinating presentation, following by an equally fascinating Q&A, culminated in a pleasant Happy Hour among professors, colleagues, and classmates.



Rating the Oral English Proficiency Test

Purdue has growing numbers of international graduate students, growing competition among graduate students for financial support through teaching assistantships, and growing pressure for high quality undergraduate instruction. All of these factors converge to some degree in the OEPP where we play an important role in determining which international graduate students are ready to accept the responsibilities of undergraduate instructional positions without English language support. The Oral English Proficiency Test is used to determine the eligibility of as many as 500 L2 English-speaking graduate students per year for teaching assistantships.

A big part of my responsibility as OEPP Testing Coordinator is the ongoing training of OEPT raters. The goal of OEPT rater training is for all raters to understand and apply the OEPT scale similarly and consistently. If the rating scale is applied to all exams similarly and consistently, OEPT scores can be considered fair for

examinees and reliable for score users. Score users here primarily refers to academic departments who ultimately decide which grad students receive TA assignments in undergraduate classrooms. Departments rely on the meaning of the OEPT test scores to determine their students' levels of oral English proficiency and appropriateness for various types of teaching positions.

The OEPT is a performance test and the examinee test responses are complex. The test is rated holistically on a six-point scale; Tests do not receive separate scores for any subset of language proficiency or any subsection of the test.

Ideally, but from a purely administrative perspective, all raters of the same test would always award it the same score. Approaching this ideal is a challenge and requires ongoing work. However, this work is where communities of practice are built – with a group of a dozen human raters using a six-point holistic scale to rate a complex language performance test. Raters bring their experiences, beliefs, and values to the rating process, and the negotiation of these experiences, values, and beliefs contributes to the development of our field, our expertise, and our understandings of the underlying constructs. We always emphasize that we train raters to rate to the scale, but training is a process of interpretation which leads to modifications of the scale itself. New voices and perspectives are incorporated into training materials and scale descriptors and our procedures develop.

An OEPT examinee performance exhibits many facets of language proficiency for raters to consider. An examinee may be strong in some respects and weak in others, may perform differently when responding to different test items. That is a very simple account of the complexity of the examinee performance side of the rating equation.

Another side of the rating equation is the rater's perception and interpretation of the examinee performance: to what degree does

the rater find the performance intelligible, comprehensible, coherent? Perception varies among raters and is influenced by many different factors such as the rater's general background, knowledge, experience and attitudes, as well as the rater's physical and mental state at the time of rating, and the rating environment.

Although the scale itself is a given, raters' understanding and interpretation of the scale may vary for many of the same reasons that raters' interpretations of examinee performances vary.

How, then, do we work toward the goal of rater agreement? Following here are the primary processes employed in the monthly OEPT rater training program this past year:

- In order to understand why we are doing what we do and why it is important, we discuss the purpose of the OEPT and the meaning and consequences of the different levels on the OEPT rating scale.
- In order to construct mutual understanding of the OEPT scale, we read about and discuss the concepts, language, and terminology of the scale descriptors. We use that language to describe test performances at different levels of the scale so that we can try to agree what the language refers to. For example, what do the concepts of 'disfluency' and 'fluency' mean in terms of language proficiency? What are different types of disfluencies, what are the varieties of fluency, and what do they sound like in different test performances by different learners who bring different language learning experiences to their performances? What is the threshold for considering a particular performance disfluent or fluent?

The Oral English Proficiency Test is used to determine the eligibility of as many as 500 L2 English-speaking graduate students per year for teaching assistantships.

Rating the Oral English Proficiency Test , Continued

- In order to construct mutual understanding of the OEPT scale, we read about and discuss the concepts, language, and terminology of the scale descriptors. We use that language to describe test performances at different levels of the scale so that we can try to agree what the language refers to. For example, what do the concepts of 'disfluency' and 'fluency' mean in terms of language proficiency? What are different types of disfluencies, what are the varieties of fluency, and what do they sound like in different test performances by different learners who bring different language learning experiences to their performances? What is the threshold for considering a particular performance disfluent or fluent?
- In order to build mutual understanding of why we assign a particular level of the scale to a test performance, we listen to benchmark exams -- exams that have been agreed upon by raters in the past to represent a particular score level -- and describe, using our rating terminology, why these tests are good examples of that score level. We do this individually in writing as well as in group discussions.
- In order to build mutual understanding of the differences between adjacent score levels, we compare

tests at a particular score level with tests at one level higher and at one level lower. We listen and try to determine and articulate what the differences are between the levels as manifested in actual performances. During group discussions, raters compare their interpretations and try to align with each other.

Rater agreement and alignment depend, in part, on raters listening carefully to each other and being willing to make adjustments to their own points of view or personal processes in order to reach agreement as a group. Absolute agreement among everyone in the group is not possible, nor is it the goal. However, score assignment must be fair to the examinees first and foremost; that is, a student's score should be a function of the performance as a reflection of a scale that has been derived through community efforts and agreement -- not a function of an individual, potentially extreme, perspective. So to be fair to the examinee and to critically examine and evaluate our process, every OEPT exam is assigned to at least two trained raters.

There are other aspects to OEPT rater training but these are the practices most central to monthly rater training. Every month OEPT raters perform these practices individually and then later as a group before the 'real' OEPT test rating begins.

OEPT raters rate 'real' tests for a few days every month and for two very intense weeks in August. Many raters rate more than a hundred OEPT exams in a year. Such exposure constitutes an apprenticeship and provides an experience that enriches the theoretical training provided in SLS/ESL classes. While rating each test conscientiously and applying the scale consistently over time is a challenge for OEPT raters, who are also ENGL 620 Instructors and graduate students, the job has many rewards. Rating generates many questions, offers opportunities for research, and provides the experience and knowledge that allows grad students to lead similar efforts when developing and administering language support programs in educational contexts outside of Purdue. Of course, there is also the knowledge, confidence, and satisfaction from knowing that we provide an important, high-quality service to the university community.

Rating OEPT exams is an important responsibility, and like many responsibilities it is both challenging and rewarding. Ever since my years as an OEPT test rater and more recently in my position as supervisor of rater training I've been fortunate to collaborate with a community of knowledgeable and hard-working grad-student raters in the OEPP. Thanks, OEPT raters, for another year of rating and for a job well done!

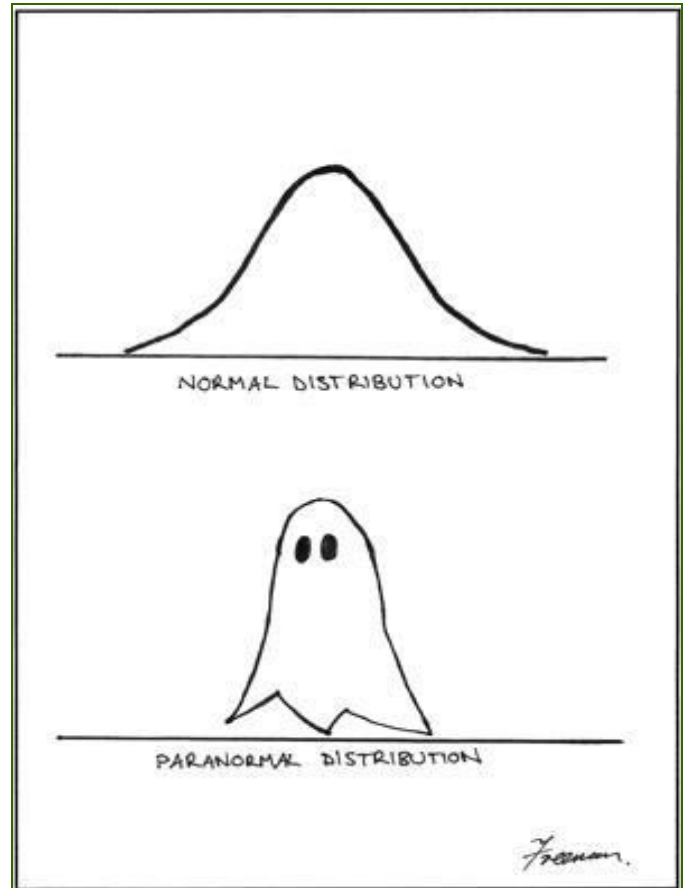
Nancy Kauper, who contributed this article, is a 2012 PhD graduate of Purdue's SLS/ESL Program. She is currently the OEPP Testing Coordinator.

Rating generates many questions, offers opportunities for research, and provides the experience and knowledge that allows grad students to lead similar efforts when developing and administering language support programs in educational contexts outside of Purdue.

Second Language Studies/ESL at Purdue is the name of a recently created group on the social networking site Facebook. The group's description reads as, "Second Language Studies/ESL at Purdue provides a platform for members of the community to connect and re-connect with current and previous classmates and colleagues. Members can share news, events, and ideas about common interests." If you have not joined, have a Facebook account, and would like to join, you are more than welcome to do so!

Have a great summer!

ESL GO wishes all SLSers a great summer! Whether you are teaching, reading, writing, researching, traveling, or whichever activity will keep you busy over the summer, enjoy your time and we will see most of you again next year! Congratulations to those who are graduating or moving on and starting a new chapter in their lives. Best wishes for a happy future!



ESL GO! Website:

<http://web.ics.purdue.edu/~eslgo/index.html>

ESL GO! Newsletter Editor

Kamal Belmihoub

Please send any suggestions, questions, or contributions to kbelmiho@purdue.edu.